Unconditional privacy over channels which cannot convey quantum information

Karol Horodecki[(1)], Michał Horodecki[(1)], Paweł Horodecki[(2)], Debbie Leung[(3)], and Jonathan Oppenheim[(4)]

[(1)]Department of Math, Physics and Computer Science, University of Gdańsk, 80–952 Gdańsk, Poland,

[(2)]Faculty of Applied Physics and Math, Technical University of Gdańsk, 80–952 Gdańsk, Poland

[(3)]Institute for Quantum Computing, University of Waterloo, Waterloo, Ontario, N2L1N8, Canada

[(4)]Deptartment of Applied Mathematics and Theoretical Physics, University of Cambridge U.K.

February 1, 2008

By sending systems in specially prepared quantum states, two parties can communicate without an eavesdropper being able to listen. The technique, called quantum cryptography, enables one to verify that the state of the quantum system has not been tampered with, and thus one can obtain privacy regardless of the power of the eavesdropper. All previous protocols relied on the ability to faithfully send quantum states. In fact, until recently, they could all be reduced to a single protocol where security is ensured though sharing maximally entangled states. Here we show this need not be the case – one can obtain verifiable privacy even through some channels which cannot be used to reliably send quantum states.

The nature of quantum systems and our ability to manipulate the state they are in has had a radical impact on the field of information theory and computation. A quantum computer can

solve problems which a classical computer cannot, and photons prepared in special states can be used to obtain privacy between two individuals sharing a fiber-optic channel – a situation impossible classically. Researchers in quantum information theory are trying to understand what aspects of quantum states and manipulations are responsible for the power of quantum computation and cryptography.

In the case of cryptography the ability to faithfully send arbitrary quantum states (*1*) appeared to lay at the heart of obtaining privacy. In the original protocol, BB84 (*2*), two-level quantum systems such as photons were faithfully sent in eigenstates of one of two complementary basis, which allows both privacy and the faithful sending of quantum states. Equivalently, entanglement based schemes (*3*) relied on the faithful distribution of maximally entangled pure states, which again allows the transmission of arbitrary states. In reality, the quantum states used in such protocols are so fragile that interaction with the environment (or the eavesdropper) causes them to rapidly decohere. However, if there is not too much noise, one can perform quantum error correction on the states, as one does in quantum computation, or post-processing on the raw key, to overcome the noise. The environment and the eavesdropper then become decoupled from the quantum states and the two parties can then obtain privacy.

Since all known protocols achieve privacy by decoupling the eavesdropper from the sent states, there was much reason to assume that this is necessary. This implied that the faithful sending of arbitrary quantum states (such as halves of maximally entangled states) appeared to be a necessary precondition for privacy. In other words, all previous cryptographic schemes are qualitatively equivalent to each other, and equivalent to distilling pure state entanglement. The first step in showing that this need not be the case was in (*4*) in the scenario where trusted states are given to the parties. There, we obtained the most general state which can produce a private key upon measurement. One can then recast all of quantum cryptography as a protocol which distills these *private states* under local operations and classical communication (LOCC). It was

then shown that there exist private states which are not equivalent to pure state entanglement. In fact, they can be produced from channels which have zero capacity (*5, 6*) – the channels cannot be used to faithfully send arbitrary quantum states, but they can produce states which are private. However, a key ingredient remained. For quantum key distribution (QKD) it is not enough for two parties to share a private state, they must be able to verify this privacy. One imagines a scenario where the eavesdropper actually gives the two parties the states, or the parties produce the states through a channel which the eavesdropper can tamper with. One must be able to verify that one indeed holds a private state and not something else.

Here, we provide a protocol which allows two parties (Alice and Bob) to verify that they indeed possess private states using only LOCC. This works for all private states, even those which can be created from zero-capacity channels, thus allowing us to obtain security over channels which cannot be used to send quantum information. The protocol is thus inequivalent to the original schemes. We previously (*7*) had introduced a protocol which worked over channels which could have arbitrary small capacity, but the protocol cannot be extended to the case where the capacity is strictly zero. Here, we will simply sketch the proof of security of our protocol. The technical details are contained in the appendix as well as in (*8*).

Let us recall that there are two scenarios for QKD. In entanglement based schemes, an adversary gives states to Alice and Bob and they distill pure entanglement in the form of the maximally entangled state $|\Phi_d\rangle := \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |i\rangle_A |i\rangle_B$ where $\{|i\rangle\}$ is a computational basis for the local systems A and B possessed by Alice and Bob respectively. They then verify that they indeed possess states very close to this form, and then measure in the computation basis to produce a secure key. One also has prepare and measure protocols, where Alice prepares a quantum state, sends it to Bob who then measures it in some basis. They then examine the results to verify that the sent states were not overly tampered with, and then perform classical post-processing on the results to obtain a key. The two schemes are equivalent in the sense that

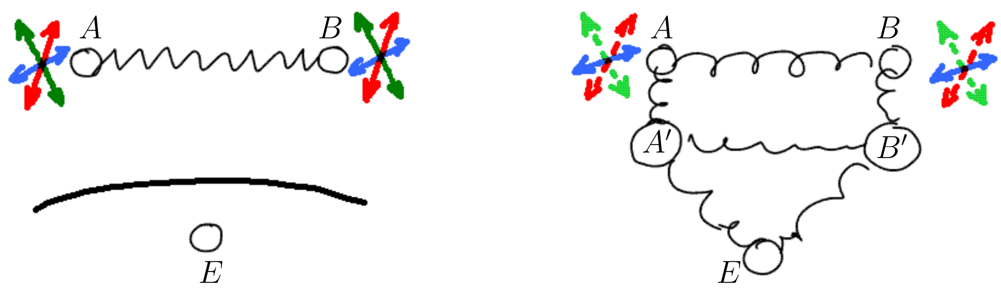

Figure 1: For a channel with non-zero capacity (left hand side), maximally entangled states can be distributed such that measuring in any direction will produce a private key. The eavesdropper is completely decoupled from the state. For the channel on the right hand side with zero capacity, the distributed states only produce a private key if the measurement is made along the blue axis. It can be shown that the eavesdropper must know at least one bit of information, but due to the ancilla on $A'B'$ (the "shield"), the bit of information is not about the key.

current prepare and measure schemes can be reduced to protocols which rely on the distillation and verification of maximally entangled states as shown in (*9*). In (*4*) it was shown that one could consider more general schemes where one considered protocols which rely on the distillation of states of the form

$$\gamma_d^U = U(|\Phi_{dAB}\rangle\langle\Phi_{dAB}| \otimes \rho_{A'B'})U^\dagger \tag{1}$$

$$U = \sum_{ij} |ij\rangle\langle ij|_{AB} \otimes U_{ijA'B'} \tag{2}$$

and viewing any protocol as the distillation and verification of such private states. Here $\rho_{A'B'}$ is an arbitrary ancilla, the $U_{ijA'B'}$ are arbitrary unitaries on it, and U is called *twisting*.

We now give the protocol for verifying private states, and prove its security. The protocol is a twisted version of verification schemes of $|\Phi_d\rangle$, and in the spirit of (*9*) we will prove security of our protocol by reducing it to security of the protocol due to Lo and Chau (*10*). Let us recall that the Lo-Chau protocol is as follows:

(1) Alice can locally prepare n systems in the state $|\Phi_2\rangle$ and distribute Bob's share to him through an untrusted channel where the eavesdropper can attack all of Bob's share at once

before it gets to him. After this step, they share the state ρ_0.

(2) Perform tests (via public but authenticated discussion) on ρ_0 by randomly selecting m_x and m_z systems, and measuring $\sigma_z \otimes \sigma_z$ to estimate the bit error rate ϵ_x, and $\sigma_x \otimes \sigma_x$ to measure the phase error rate ϵ_z. Here, the σ are the standard Pauli matrices. The error rates essentially tell us how far ρ_0 deviates from a maximally entangled state.

(3) Based on the results of the test, the parties perform an appropriate entanglement purification protocol (EPP) to ρ_0 and output a state $\tilde{\gamma}$ which will be close to the maximally entangled state with high probability. One doesn't need to know the exact form of ρ_0, but only the error rates.

(4) Generate a key by measuring $\tilde{\gamma}$ locally. The key can have varying size (depends on the error rate), and zero key-length means "abort QKD."

The security of this protocol rests on the fact that the estimates ϵ_x, ϵ_z of the two error rates by random sampling will converge with high probability to their expectation values over the entire initial state ρ_0 thus ensuring that the final state $\tilde{\gamma}$ is close to maximally entangled. For small δ and $m_z < (\frac{2\delta^2}{1+2\delta^2})n$, we have for example (*10*)

$$\Pr(\,|\epsilon_{zP} - \epsilon_z| \geq \delta) \leq 2e^{-2m_z\delta^2} \tag{3}$$

where ϵ_{zP} is the expectation value of the phase error rate. This result is from sampling theory and can be found as Proposition 1 in the appendix.

We now wish to modify this protocol so that we can use it to verify private states, which for the moment we take to be many copies of γ_2^U. In (*4, 11, 12*) examples of such states were given which result from zero-capacity channels (i.e. they are *bound entangled* (*13*)), and thus our protocol will work over such channels.

Since private states are twisted maximally entangled states, we could achieve verifiable privacy, by *untwisting* the private state before each step of the protocol, so that we are just

acting the above protocol on the maximally entangled state. We would thus need to modify the protocol as follows:

(2') Apply untwisting $U^{\otimes n\dagger}$ to ρ_0, then estimate ϵ_x and ϵ_z on the $(AB)^{\otimes n}$ systems as in the original step (2), and finally reapply $U^{\otimes n}$.

(3') Apply untwisting $U^{\otimes n\dagger}$, measure out a "raw-key" in the computational basis of the remaining $n-m_x-m_z$ systems.

(4') Perform error correction and privacy amplification on the raw-key via public discussion.

Such a protocol is unfeasible since U may be a global unitary and cannot be done using only LOCC. However, it is secure, since if we were able to perform the twisting and untwisting, the only difference between this protocol and that of the Lo-Chau one is that classical privacy amplification (*14*) and error correction is used instead of the entanglement purification protocol (EPP). This does not effect security, since it was shown (*15,16*) that there exist classes of EPPs such that applying the EPP and measuring out a key can be securely converted to protocols where a key is first measured out and then we apply classical error correction and privacy amplification on the raw key to obtain a secure one. We now explain how to convert the above unfeasible protocol to a feasible one which can be performed via LOCC.

First, in step (2'), for the $n-m_x-m_z$ which are not used for testing, the twisting and untwisting cancel and therefore, do not need to be performed. Also, the measurement of bit errors via $\sigma_z \otimes \sigma_z$ on AB commutes with the twisting and untwisting, and therefore, the twisting and untwisting cancel. Similarly, in step (3'), the measurement commutes with the untwisting, and therefore this untwisting is also unnecessary. Finally, for step (2'), untwisting the state, estimating the expected number of phase errors, and retwisting is equivalent to estimating the twisted phase error rate via the operator $\Sigma_x = U_{ABA'B'}(\sigma_x \otimes \sigma_x \otimes I_{A'B'})U^{\dagger}_{ABA'B'}$. Mercifully, our only remaining task is to find a way to estimate this error rate via LOCC, rather than via direct measuring of the global operator Γ_x.

To do this, we will first decompose Σ_x in terms of products of observables which can be locally measured by Alice and Bob. We then show that this estimation of the observable in terms of these product observables is a good estimation. As will be explained shortly, this involves adapting the quantum deFinetti theorem (*17*) and a Chernoff-like bound.

We can always decompose any observable into product observables. In particular:

$$\Sigma_x = U_{ABA'B'}(\sigma_x \otimes \sigma_x \otimes I)U^\dagger_{ABA'B'} \tag{4}$$

$$= \sum_{j_a,j_b=1}^{t} s_{j_a j_b} O_{j_a AA'} \otimes O_{j_b BB'} \tag{5}$$

where $\{O_j\}_{j=1}^t$ is a basis (trace-orthonormal) for hermitian operators acting on AA' and BB', and $t = d^2 d'$. Alice and Bob can now estimate the average value of Σ_x by dividing the m_z samples into t^2 groups, and then estimating individually $O_{j_a AA'}$ and $O_{j_b BB'}$ on the ith test system. They then multiply their results publicly, and finally sum these products over $i = 1, \cdots, m_z/t^2$ with the coefficients given by Eq. (5).

The outcome of this LOCC estimation procedure will result in giving some emperical value for the average of Σ_x, which we call $\langle \Sigma^{ind} \rangle_{emp}$. We want to compare $\langle \Sigma^{ind} \rangle_{emp}$ to the emperical value $\langle \Sigma \rangle_{emp}$ obtained from estimating Σ_x via a direct global measurement (which is the measurement that is performed in the unfeasible yet secure modified Lo-Chau protocol). If the two values are close, then we have shown that the LOCC measurement is a good estimation of Σ_x

Indeed $\langle \Sigma^{ind} \rangle_{emp}$ will be close to $\langle \Sigma \rangle_{emp}$ if the entire m_z sample systems are in a joint tensor-power state $\rho_0^{\otimes n}$, and if the number of systems we test is large enough. This follows from Eq. (5) and the fact that for tensor power states, we may regard each measurement as an independent event. We can then use the Chernoff bound which states that a random sample of k independent measurements of an operator O on state $\rho^{\otimes n}$ will converge exponentially fast in k to its average value $\bar{O} = \mathrm{Tr}(O\rho)$. More precisely, the probability that $|O - \bar{O}| \geq \delta$ decays as $\sim e^{-Ck\delta^2}$ for C a positive constant. In this case we know that the estimate of each of the t^2

local measurements will converge exponentially fast to $\mathrm{Tr}(\rho_0 O_j)$ as we increase the number of tested systems $k = m_z/t^2$.

However, in our current problem, Alice and Bob share ρ_0 which is *not* a tensor-power state, and each measurement cannot be considered to be an independent event. Fortunately, there is a sense in which a random sampling of m_z systems is close to tensor-power. First, permutation symmetry can be imposed on the protocol (since we can choose a random sample in any order), and second, since the estimation involves only a small portion (m_z) of the entire n systems, the exponential quantum deFinetti theorem (*17*) states that the measured (reduced) state is close to a mixture of "almost-tensor-power-states". This is captured by Theorem 2 of the appendix. We can now apply a Chernoff-like bound to these almost-tensor-power-states. The exact analysis involves many adaptations of the results in (*17*) and is given in the appendix as Theorem 1. The result has consequences well beyond the current considerations. Essentially, any realizations of an observable (i.e. a decomposition of the operator in terms of others), is a good one, in the sense that performing one kind of measurement on m out of n systems via one realization of the measurement, will yield average values which are well correlated with the values obtained by performing another realization of the measurement on the remaining $n - m$ systems. This is captured in Theorem 3 of the appendix. We can apply this to the current case to show that the probability that $|\langle \Sigma^{ind} \rangle_{emp} - \langle \Sigma \rangle_{emp}| > \delta$ can be made small. This says that the estimated twisted phase errors through measuring a sample via LOCC is correlated with the result we would obtain if we made an ideal measurement of twisted phase errors on the rest of system. Thus in terms of security, the only difference between the modified protocol, and that of Lo-Chau, is that instead of Equation 3 governing the accuracy of the phase error estimate, we have

through Theorem 3

$$
\begin{aligned}
& \Pr |\langle \Sigma^{ind} \rangle_{emp}^{(m)} - \langle \Sigma \rangle_{emp}^{(m+n)}| > \delta \\
\leq \quad & 2e^{-\frac{(n-m_z)(r+1)}{2n} + \frac{1}{2} d^4 d'^2 \ln(n-m_z)} \\
+ \quad & (t^2+1) 2^{-\left[\frac{\delta^2}{36t^2 d^2 d'} - H(\frac{rt^2}{m_z})\right] \frac{m_z}{t^2} + d' d^2 \log(\frac{m_z}{2t^2}+1)} \\
+ \quad & d' d^2 e^{-\frac{m_z \delta^2}{72 d'^2 d^4 t^2}} \qquad (6)
\end{aligned}
$$

where the three expressions in the upper bound respectively come from the exponential quantum deFinetti theorem, the Chernoff bound, and random sampling theory. Here, d is the dimension of the maximally entangled state and we can take $d = 2$, d' the dimension of each ancilla on $A'B'$, and r is some natural number we will take to be $\geq d^4 d'^2 \ln n$. The superscripts for the emperical values of Σ_x refer to $\langle \Sigma^{ind} \rangle_{emp}$ being measured using m systems while $\langle \Sigma \rangle_{emp}$ is measured on the remaining $n - m$.

This then proves security of the entire scheme, since the only significant change from the unfeasible modified protocol is a different method for estimating phase errors. The calculation of security in terms of composable security parameters for QKD (*18*) is given in (*8*).

We now touch on several issues which arise. The protocol we have given, as with all entanglement based protocols, relies on keeping the quantum state ρ_0 from decohering throughout the procedure, and it is therefor not currently practical. However, it can be converted to a prepare-and-measure protocol where Alice prepares a state, sends it down a channel (which might have zero quantum capacity), and then Bob measures the state right away. The conversion adapts well known techniques and is contained in (*8*) along with an example.

Next, in the above protocol, we considered verification of tensor powers of private states with dimension two on A i.e. $\gamma_2^{\otimes n}$ under general attacks. It is straightforward to extend this to the verification of private states of any dimension, and states where the twisting is close to tensor power. It is not clear whether one can extend this to private states which are not tensor

power such as a single γ_d; as of yet we do not have a no-go theorem. This is quite different from verification of pure state entanglement where the maximally entangled state of any dimension can be written as $|\Phi_2\rangle^{\otimes n}$ and we are thus always trying to verify something close to tensor power.

Here, we considered a twisted version of the Lo-Chau scheme, but we could have just as well considered twisted versions of other parameter estimation schemes. Indeed our protocol is not optimal in its use of resources and it may be interesting to improve it. Some potential avenues were noted in (*8*). A tomographic verification scheme was suggested originally in (*4*), and it may be interesting to explore its efficiency. It is simpler in the sense that one could just discard some states, and be left with almost-tensor-product states as in (*17*).

Finally, here we have demonstrated conceptually that quantum key distribution is not equivalent to the ability to send quantum information. However, we only know of a few channels and set of states which have the property of offering security without allowing quantum communication. It would be very interesting to find other examples, and perhaps even more interesting to know whether there are any bound entangled states (and the corresponding zero-capacity channels) which cannot produce a secure key.

Acknowledgments We thank Daniel Gottesman and Hoi-Kwong Lo for valuable discussions. We acknowledge support from EU grant QAP IST-015848 and IP SCALA 015714. JO also acknowledges the Royal Society and DL is supported by the CRC, CRC-CFI, ORF, CIAR, NSERC, MITACS, and ARO. KH acknowledges the support of the Foundation for Polish Science.

Appendix

In Section 1 we present Theorem 1 on the extent to which a permutationally invariant set of systems behaves like independently and identically distributed (IID) states for the purpose of parameter estimation. This is an application of the quantum de Finetti theorem and the generalized Chernoff bound. Section 1.2 presents some results which will be used in Section 1.3. It is in this latter section where the key Theorem 3 is presented in Subsection 1.3.3. It relates the distance between direct measurements and indirect measurements of an observable.

1 LOCC estimation of the expectation of an IID observable

1.1 Finite quantum de Finetti theorem and generalized Chernoff bound

We say that a state ρ_n on Hilbert space $\mathcal{H}^{\otimes n}$ satisfies the Chernoff bound with respect to a state σ on $\mathcal{H}$ and a measurement $\mathcal{M}$ on $\mathcal{H}$ if (with high probability) the *frequency distribution* obtained by measuring $\mathcal{M}^{\otimes n}$ on ρ_n is close to that of measuring $\mathcal{M}$ on σ. For example, $\rho_n = \sigma^{\otimes n}$. However many other states satisfy the same property. An important class is called *almost power states*, which are formulated and studied in (*17*). We adapt results in (*17*) for our own purpose in the following.

Theorem 1 (Finite quantum de Finetti theorem plus Chernoff bound) *Consider any permutationally invariant (possibly mixed) state ρ_{n+k} on Hilbert space $\mathcal{H}^{\otimes(n+k)}$. Let $\rho_n = \mathrm{Tr}_k \rho_{n+k}$ be the partial trace of ρ_{n+k} over k systems. Then there exists a probability measure μ on (possibly mixed) states σ acting on $\mathcal{H}$ and a family of states $\rho_{n,r}^{(\sigma)}$ such that*

1. *The state ρ_n is close to a mixture of the states $\rho_{n,r}^{(\sigma)}$*

$$\left\| \rho_n - \int \rho_{n,r}^{(\sigma)} \, d\mu(\sigma) \right\|_{\mathrm{tr}} \leq 2\, e^{-\frac{k(r+1)}{2(n+k)} + \frac{1}{2}\dim(\mathcal{H})^2 \ln k} \tag{7}$$

2. *The states* $\rho_{n,r}^{(\sigma)}$ *(called* almost power states*) satisfy the Chernoff bound in the following sense*

$$\Pr\left(\left\|P_{\mathcal{M}}(\sigma) - Q_{\mathcal{M}}(\rho_{n,r}^{(\sigma)})\right\| > \delta\right) \\ \leq 2^{-n\left[\frac{\delta^2}{4} - h(\frac{r}{n})\right] + |W|\log(\frac{n}{2}+1)} =: e(\delta) \tag{8}$$

where $\mathcal{M}=\{M_w\}_{w\in W}$ *is any measurement on* $\mathcal{H}$, $P_{\mathcal{M}}(\sigma) = \{\mathrm{Tr}(\sigma M_w)\}_w$, $Q_{\mathcal{M}}(\rho_{n,r}^{(\sigma)})$ *is the frequency distribution obtained from measuring* $\mathcal{M}^{\otimes n}$ *on the state* $\rho_{n,r}^{(\sigma)}$*, and* $|W|$ *is the size of the alphabet* W*.*

3. *Reduced density matrices of the states* $\rho_{n,r}^{(\sigma)}$ *(to* $n' \leq n$ *systems) satisfy the same Chernoff bound:*

$$\Pr(\left\|P_{\mathcal{M}}(\sigma) - Q_{\mathcal{M}}(\rho_{n,r,n'}^{(\sigma)})\right\| > \delta) \\ \leq 2^{-n'\left[\frac{\delta^2}{4} - h(\frac{r}{n'})\right] + |W|\log(\frac{n'}{2}+1)} \tag{9}$$

where $r \leq n'/2$ *and* $\rho_{n,r,n'}^{(\sigma)} = \mathrm{Tr}_{n-n'}\rho_{n,r}^{(\sigma)}$ *is the partial trace of* $\rho_{n,r}^{(\sigma)}$ *over* $n - n'$ *systems.*

Proof: We first collect various facts, definitions, and results from (*17*).

1.1.1 Facts and definitions

Definition 1 Almost power state: (Def. 4.1.4, in (*17*)) *Suppose* $0 \leq r \leq n$*. Let* $\mathrm{Sym}(\mathcal{H}^{\otimes n})$ *denote the symmetric subspace of pure states of Hilbert space* $\mathcal{H}^{\otimes n}$*. Let* $|\theta\rangle \in \mathcal{H}$ *be an arbitrary pure state and consider:*

$$\mathcal{V}(\mathcal{H}^{\otimes n}, |\theta\rangle^{\otimes n-r}) := \{\pi(|\theta\rangle^{\otimes n-r} \otimes |\psi_r\rangle) : \pi \in S_n, \\ |\psi_r\rangle \in \mathcal{H}^{\otimes r}\}$$

where S_n *is the permutation group of the* n *systems. We define the almost power states along* $|\theta\rangle$ *to be the set of pure states in*

$$|\theta\rangle^{[\otimes,n,r]} := \mathrm{Sym}(\mathcal{H}^{\otimes n}) \cap \mathrm{span}(\mathcal{V}(\mathcal{H}^{\otimes n}, |\theta\rangle^{\otimes n-r})) \tag{10}$$

We denote the set of mixtures of almost tensor power states along $|\theta\rangle$ *as* $\mathrm{conv}(|\theta\rangle^{[\otimes,n,r]})$.

With the above definition, we shall prove the following lemma:

Lemma 1 *If* $\varrho_n \in \mathrm{conv}(|\theta\rangle^{[\otimes,n,r]})$*, then,* $\varrho_{n-m} \in \mathrm{conv}(|\theta\rangle^{[\otimes,n-m,r]})$ *where* $\varrho_{n-m} = \mathrm{Tr}_m(\varrho_n)$ *is the reduced density matrix after the partial trace over any* m *out of the* n *systems (by symmetry, without loss of generality, we take the first* m *systems).*

Proof .-

Since membership in $\mathrm{conv}(|\theta\rangle^{[\otimes,n-m,r]})$ is preserved under mixing, it suffices to prove the lemma for pure $\varrho_n = |\Psi\rangle\langle\Psi|$, with $|\Psi\rangle \in |\theta\rangle^{[\otimes,n,r]}$.

We can pick an ensemble realizing ϱ_{n-m} of our choice, and prove the lemma by showing that any element $|\Psi_{n-m}\rangle$ in that ensemble belongs to $|\theta\rangle^{[\otimes,n-m,r]}$. Our ensemble is obtained by an explicit partial trace of $|\Psi_n\rangle$ over the first m subsystems along the computational basis. An element is given by

$$|\Psi_{n-m}\rangle = \langle i_1|...\langle i_m| \otimes I_{n-m}|\Psi_n\rangle. \tag{11}$$

Now, we note two facts:

(i) $|\Psi_{n-m}\rangle \in \mathrm{Sym}(\mathcal{H}^{\otimes(n-m)})$ – This is because $|\Psi_n\rangle \in \mathrm{Sym}(\mathcal{H}^{\otimes(n)}) = \mathrm{span}(|\phi\rangle^{\otimes n})$.

(ii) $|\Psi_{n-m}\rangle \in \mathcal{V}(\mathcal{H}^{\otimes n-m}, |\theta\rangle^{\otimes(n-m-r)})$ – This is because $|\Psi_n\rangle \in \mathcal{V}(\mathcal{H}^{\otimes n}, |\theta\rangle^{\otimes n-r})$, and expressing $|\Psi_n\rangle$ in terms of the spanning vectors of $\mathcal{V}(\mathcal{H}^{\otimes n}, |\theta\rangle^{\otimes n-r})$ and putting it into Eq. (11), we have

$$|\Psi_{n-m}\rangle = \sum_{\Psi_r,\pi} \alpha_{\Psi_r,\pi}\langle i_1|\cdots\langle i_m| \otimes I_{n-m}\ \pi\left(|\theta\rangle^{\otimes n-r} \otimes |\Psi_r\rangle\right).$$

Elementary analysis shows that any term of the above sum is, up to permutation, of the form $(\langle i_1|\theta\rangle)\cdots(\langle i_p|\theta\rangle)|\theta\rangle^{\otimes n-r-p} \otimes [\langle i_{p+1}|\cdots\langle i_m| \otimes I_{r-(m-p)}\ \pi'(|\Psi_r\rangle)]$ where $0 \leq p \leq m$, and "absorbing" $m-p$ copies of θ to the last part of the vector, we get $|\theta\rangle^{\otimes n-(r+m)} \otimes |\Psi_r''\rangle$. Thus, $|\Psi_{m-n}\rangle$ is a sum of terms of the form $\pi(|\theta\rangle^{\otimes n-(r+m)} \otimes |\Psi_r''\rangle)$, and belongs to $\mathcal{V}(\mathcal{H}^{\otimes(n-m)}, |\theta\rangle^{\otimes n-(r+m)})$. This proves the second fact, and also the lemma. □

Property of a mixture of almost tensor power states behaves approximately like a mixture of tensor power states, so that the generalized version of Chernoff bound holds.

Lemma 2 (Theorem 4.5.2 of (*17*)) *Let* $\mathcal{M} = \{M_w\}_{w\in\mathcal{W}}$ *be a POVM on* $\mathcal{H}$*, let* $0 \le r \le \frac{n}{2}$*. Moreover let* $|\theta\rangle \in \mathcal{H}$ *and let* $|\Psi_n\rangle$ *be a vector from* $|\theta\rangle^{[\otimes,n,r]}$*. There holds:*

$$\begin{aligned} &P(||P_{\mathcal{M}}(|\theta\rangle\langle\theta|) - P_{\mathcal{M}}[|\Psi_n\rangle\langle\Psi_n|]|| > \delta) \\ &\quad \le 2^{-n\left[\frac{\delta^2}{4} - h(\frac{r}{n})\right] + |W|\log(\frac{n}{2}+1)} =: e(\delta) \end{aligned}$$

where $P_{\mathcal{M}}(|\theta\rangle\langle\theta|) = \{\mathrm{Tr}|\theta\rangle\langle\theta|M_w\}_w$ *and* $P_{\mathcal{M}}[|\Psi_n\rangle\langle\Psi_n|]$ *is the frequency distribution of outcomes of* $\mathcal{M}^{\otimes n}$ *applied to* $|\Psi_n\rangle\langle\Psi_n|$*, and the probability is taken over those outcomes. Note that we have used* $e(\delta)$ *instead of* $\delta(e)$ *in (17).*

Consider the general probability $\Pr(\|P_{\mathcal{M}}(\rho) - P_{\mathcal{M}}[\varrho_n]\| < \delta)$ where $P_{\mathcal{M}}[\varrho_n]$ is a frequency distribution of outcomes of $\mathcal{M}^{\otimes n}$ applied to $|\Psi_n\rangle\langle\Psi_n|$. The distribution $P_{\mathcal{M}}[\varrho_n]$, if treated as a functional of ϱ_n on the space $\mathcal{H}^{\otimes n}$, is *linear* in ϱ_n. Following this we get immediately:

Corollary 1 *Lemma 2 holds when replacing the projector* $|\Psi_n\rangle\langle\Psi_n|$ *(for* $|\Psi_n\rangle \in |\theta\rangle^{[\otimes,n,r]}$*) by* $\varrho_n \in \mathrm{conv}(|\theta\rangle^{[\otimes,n,r]})$.

Apart form the generalised Chernoff-type lemmas, we also need the crucial exponential quantum finite deFinetti theorem:

Theorem 2 (Theorem 4.3.2 of (*17*)) *For any pure state* $|\psi_{n+k}\rangle \in \mathrm{Sym}(\mathcal{H}^{\otimes n+k})$ *and* $0 \le r \le n$ *there exists a measure* $d\nu(|\theta\rangle)$ *on* $\mathcal{H}$ *and for each* $|\theta\rangle \in \mathcal{H}$ *a pure state* $|\psi_n^{(\theta)}\rangle \in |\theta\rangle^{[\otimes,n,r]}$ *such that*

$$\begin{aligned} &\left\|\mathrm{Tr}_k|\psi_{n+k}\rangle\langle\psi_{n+k}| - \int_{\mathcal{H}} |\psi_n^{(\theta)}\rangle\langle\psi_n^{(\theta)}|d\nu(|\theta\rangle)\right\|_{\mathrm{tr}} \\ \le\ & 2e^{-\frac{k(r+1)}{2(n+k)} + \frac{1}{2}\dim(\mathcal{H})\ln k} \end{aligned} \tag{12}$$

Finally, we need the fact that any permutationally invariant state has a symmetric purification.

Lemma 3 (Lemma 4.2.2 of (*17*)) *Let ρ_n be permutationally invariant state on $\mathcal{H}$. Then there exists purification of the state on* $\mathrm{Sym}((\mathcal{H}\otimes\mathcal{H})^{\otimes n})$

This concludes the list of facts and definitions needed for proving Theorem 1.

1.1.2 Proof of Theorem 1

Consider an arbitrary permutationally invariant state ϱ_{n+k} on Hilbert space $\mathcal{H}^{\otimes(n+k)}$.

Step (1): According to Lemma 3 there is a purification $|\psi_{n+k}\rangle$ that belongs to $\mathrm{Sym}(\mathcal{H}'^{\otimes n+k})$ where $\mathcal{H}'=\mathcal{H}\otimes\tilde{\mathcal{H}}$ and dim$(\tilde{\mathcal{H}})$ =dim$(\mathcal{H})$.

Step (2): We apply to ψ_{n+k} theorem 2 with the changes

$$\begin{aligned} \mathcal{H} &\rightarrow \mathcal{H}'=\mathcal{H}\otimes\tilde{\mathcal{H}} \\ d &\rightarrow d^2 \end{aligned} \tag{13}$$

Step (3): After application of theorem 2 we perform partial trace over $\tilde{\mathcal{H}}^{\otimes n}$, the purifying spaces introduced in (1). We denote this partial trace by $\tilde{\mathrm{Tr}}$. This partial trace induces from the measure on $\mathcal{H}'$ in step (2) the new measure $\mu(\sigma)$ on the set of all mixed states σ defined on $\mathcal{H}$. (This is defined by probability ascribed by the measure μ to the subset of $\mathcal{H}'$ equal to the equivalence class of all pure states $|\theta\rangle$ which satisfy $\tilde{\mathrm{Tr}}(|\theta\rangle\langle\theta|)=\sigma$). This partial trace produces also the states $\rho_{n,r}^{(\sigma)}$ defined directly by $\rho_{n,r}^{(\sigma)}\equiv\tilde{\mathrm{Tr}}(|\psi_n^{(\theta)}\rangle\langle\psi_n^{(\theta)}|)$ where the existence of the pure states $|\psi_n^{(\theta)}\rangle$ is guaranteed by theorem 2. Finally we note that partial trace does not increase the trace distance between two quantum states, so applying partial trace to the LHS of (12) and using the notation described above we get immediately the inequality (7). This proves the first item of Theorem (1).

To prove the second item of Theorem (1), remember from the above that $\rho_{n,r}^{(\sigma)}\equiv\tilde{\mathrm{Tr}}(|\psi_n^{(\theta)}\rangle\langle\psi_n^{(\theta)}|)$. Since $|\psi_n^{(\theta)}\rangle$ is an almost power pure state, lemma 2 applies. Further, it holds for all POVM-s on

$\mathcal{H}' = \mathcal{H} \otimes \tilde{\mathcal{H}}$, and in particular for incomplete POVM-s acting only on $\mathcal{H}$ but not on $\tilde{\mathcal{H}}$. Thus, the conclusion of lemma 2 holds with the change: $\mathcal{M} \rightarrow \mathcal{M} \otimes I$, which gives item (2).

Finally, to prove item 3 of theorem 1, note that the reduced density matrices $\varrho^{\sigma}_{n,r,n'}$ of interest can be obtained from the pure state $|\psi_n^{(\theta)}\rangle$ above by tracing (i) first over $n - n'$ subsystems corresponding to $\mathcal{H}'$, producing a state on $\mathcal{H}'^{\otimes n'}$, and (ii) then over n' subsystems corresponding to $\tilde{\mathcal{H}}$.

Then lemma 1 guarantees that the first partial trace produces a mixed state $\varrho_{n'}$ in $\mathrm{conv}(|\theta\rangle^{[\otimes,n',n'-r]})$ (with underlying space $\mathcal{H}'$. Applying corollary 1 to $\varrho_{n'}$ with n' instead of n, it suffices to consider a pure state in $|\theta\rangle^{[\otimes,n',n'-r]}$. Finally, lemma 2 can be applied to this pure state with $\mathcal{M} \rightarrow \mathcal{M} \otimes I$ which concludes item 3 (with the assumption $0 \leq r \leq \frac{n'}{2}$). $\square$

1.2 Two other useful results

1.2.1 Classical random sampling

In addition to the fact and definitions above and Theorem 1, we will need the following result on classical random sampling (or equivalently symmetric probability distribution).

Proposition 1 (Classical sampling theory) Lemma A.4 from (*19*). *Let Z be an n-tuple and Z' a k-tuple of random variables over set $\mathcal{Z}$, with symmetric joint probability P. Let $Q_{z'}$ be a frequency distribution of a fixed sequence z' and $Q_{(z,z')}$ be frequency distribution of a sequence (z, z'). Then for every $\epsilon \geq 0$ we have*

$$P(||Q_{(z,z')} - Q_{z'}|| \geq \epsilon) \leq |\mathcal{Z}| e^{-k\epsilon^2/8|\mathcal{Z}|}. \tag{14}$$

The result says that frequency obtained from a small sample is close to frequency distribution obtained from the whole system.

1.2.2 From probabilities to averages

Lemma 4 *Consider an observable L on Hilbert space $\mathcal{H}$, $\dim \mathcal{H} = d$. Let $L = \sum_{i=1}^{t} s_i L_i$, where L_i satisfy $\mathrm{Tr} L_i L_j^\dagger = \delta_{ij}$. Let eigenvalues of L_i be denoted by $\lambda_l^{(i)}$. Consider arbitrary state ρ, and let $P^{(i)} = \{p_l^{(i)}\}$ be the probability distribution on l (which eigenvalue) induced by measuring L_i on ρ. Let $Q^{(i)} = \{q_l^{(i)}\}$ be an arbitrary family of distributions on eigenvalues of L_i. We then have*

$$\begin{aligned} & |\langle L\rangle_\rho - \sum_i s_i \sum_l \lambda_l^{(i)} q_l^{(i)}| \\ \leq \quad & \sqrt{t} ||L||_{HS} \max_i ||P^{(i)} - Q^{(i)}||, \end{aligned} \tag{15}$$

where $\|\cdot\|_{HS}$ is the Hilbert-Schmidt norm, $\|\cdot\|_\infty$ is the operator norm, and $\|\cdot\|$ is the trace norm.

Proof

$$\begin{aligned} & \left| \langle L\rangle_\rho - \sum_i s_i \sum_l \lambda_l^{(i)} q_l^{(i)} \right| \\ = \quad & \left| \sum_i s_i \sum_l \lambda_l^{(i)} (p_l^{(i)} - q_l^{(i)}) \right| \\ \leq \quad & \sum_i |s_i| \, (\max_l |\lambda_l^{(i)}|) \, \|P^{(i)} - Q^{(i)}\| \\ = \quad & \sum_i s_i \|L_i\|_\infty \, \|P^{(i)} - Q^{(i)}\| \\ \leq \quad & (\max_j \|P^{(j)} - Q^{(j)}\|) \sum_i s_i \, \|L_i\|_\infty \end{aligned} \tag{16}$$

Since $||L_i||_\infty = 1$, using convexity of x^2 we obtain

$$\sum_{i=1}^{t} s_i \, \|L_i\|_\infty = \sum_i s_i \leq \sqrt{t} \sqrt{\sum_i s_i^2} = \sqrt{t} \, \|L\|_{HS} \tag{17}$$

which ends the proof. □

1.3 Estimation - detailed description

We consider $2m+n$ systems with Hilbert space $\mathcal{H}^{\otimes(2m+n)}$, $\dim\mathcal{H} = d$ in a permutationally invariant state ϱ_{2m+n}. Suppose the ultimate goal is to obtain the "frequency mean-value" of some single-system observable Σ on $n+m$ systems. In other words, we want to measure $\frac{1}{N}\sum_{j=1}^{N}\Sigma^{(j)}$ where $\Sigma^{(j)} = I\otimes I\otimes\cdots\otimes\Sigma\otimes\cdots\otimes I$ on the N subsystems for $N = n+m$.

Because of experimental limitations (here, it is the LOCC constraints on Alice and Bob), they are restricted to measuring product operators of the form $L = L_A\otimes L_B$ by independently finding the eigenvalues of L_A and L_B (i.e., making the measurements $L_A\otimes I$ and $I\otimes L_B$), discussing over classical channels and multiplying their outcomes together. Now, to measure Σ, one can first rewrite it in terms of product operators L_i:

$$\Sigma = \sum_{i=1}^{t} s_i L_i \tag{18}$$

where we have chosen $\{L_i\}$ to be hermitian and trace orthonormal, so that s_i are real. The L_i-s are "*intermediate observables*." We will describe an inference scheme that (1) involves only the estimation of the "frequency mean-value" of Σ on a small number (m) of subsystems, and (2) the measurement of Σ is done indirectly via measurements of the L_i's.

The analysis will start with special assumption about the $2m$-element sample, m of which are used for indirect estimation. The assumptions are relaxed on that sample. After that properties of the other $m+n$ subsystems are inferred.

1.3.1 Analysis of the $2m$ sample in an "almost power state along σ":$\varrho^{(\sigma)}_{2m,r}$

Suppose the first $2m$ subsystems are in a joint state $\varrho^{(\sigma)}_{2m,r}$, with $r\le\frac{1}{2}\times 2m$. We expect the state $\varrho^{(\sigma)}_{2m,r}$ to play a role similar to the state $\sigma^{\otimes 2m}$. Define the theoretical direct average

$$\langle\Sigma\rangle_\sigma = \mathrm{Tr}(\Sigma\sigma) = \sum_i s_i\langle L_i\rangle_\sigma \tag{19}$$

We will show that the empirical average, either obtained directly or indirectly, will be close to the above.

For the indirect measurement, divide the first m subsystems into t groups. Each group has $m' = m/t$ subsystems. Alice and Bob take the ith group ($i = 1, \cdots, t$) and measure L_i on each site as described above (the measurement is $\mathcal{L}_i$). In other words, the measurement $M^{\mathrm{indirect}} = \otimes_{i=1}^{t}(\mathcal{L}_i^{\otimes m'})$ is applied to the first m subsystems of the entire $2m + n$ subsystems. The reduction of the state $\varrho_{2m,r}^{(\sigma)}$ to the first m subsystems induces probability distribution $\mathcal{P}$ on the outcomes of M^{indirect}.

Since we expect $\varrho_{2m,r}^{(\sigma)}$ and $\sigma^{\otimes m}$ to behave similarly, consider the probability distribution on alphabet $\mathcal{A}_i$ of observable $L_i = \sum_l \lambda_l^i \overline{P}_l^{(i)}$ induced by the state σ as follows:

$$P_i = \{\mathrm{Tr}(\sigma \overline{P}_l)\}_l \tag{20}$$

An execution of the measurement $\mathcal{L}_i^{\otimes m'}$ gives a particular outcome $(l_1, ..., l_{m'})$ and induces frequency distribution Q_i on alphabet $\mathcal{A}_i$ of the observable L_i.

Then, the empirical frequency distributions Q_i is close to the "theoretical" distribution P_i:

Fact 1

$$\mathcal{P}(\|P_i - Q_i\| \geq \delta) \leq e(\delta, m', r, d), \tag{21}$$

where d is the dimension of the single site Hilbert space, and

$$e(\delta, n, r, |Z|) := 2^{-(\frac{\delta^2}{4} - H(\frac{r}{n}))n + |Z| \log(\frac{n}{2}+1)} \tag{22}$$

Proof - Follows immediately from the third item of Theorem 1. Note that we use item (3) not (2) since we perform the measurement only on *part* of the state $\varrho_{2m,r}^{(\sigma)}$.

Remark - Note also that P_i is constant while Q_i is a random variable.

Now, we define the theoretical average values for the intermediate observables L_i's:

$$\langle L_i \rangle_\sigma = \mathrm{Tr}(L_i \sigma) \tag{23}$$

and the empirical average

$$\langle L_i\rangle_{emp} = \sum_l \lambda_l^{(i)} Q_i(l) \tag{24}$$

where $Q_i(l)$ denotes value of Q_i on specific event l from alphabet $\mathcal{A}_i$. (Again, $\langle L_i\rangle_\sigma$ is constant while $\langle L_i\rangle_{emp}$ is a random variable depending on the particular outcomes of measurement - recall that $L_i = \sum_i \lambda_l^{(i)} \bar{P}_l^{(i)}$). And again, recall that we the *empirical* value of Σ obtained *indirectly*, *via* empirical distributions of the L_i.

$$\langle \Sigma^{ind}\rangle_{emp}^{(m)} = \sum_i s_i \langle L_i\rangle_{emp} \,. \tag{25}$$

We now show that the indirect empirical average is close to the direct theoretical average in Eq. (19). First applying the union bound to Fact 1, we get

$$\mathcal{P}(\cup_{i=1,\dots,t}\{||P_i - Q_i|| > \delta\}) \leq t\cdot e(\delta, m', r, d) \tag{26}$$

Then using Lemma 4 we obtain that $P(|\sum_i^t s_i\langle L_i\rangle_\sigma - \sum_i^t s_i \langle L_i\rangle_{emp}| > \delta) \leq t\cdot e(\frac{\delta}{||\Sigma||_{HS}\sqrt{t}}, m', r, d)$ which is just

$$\mathcal{P}(|\langle\Sigma\rangle_\sigma - \langle\Sigma^{ind}\rangle_{emp}^{(m)}| > \delta) \leq t\cdot e(\frac{\delta}{||\Sigma||_{HS}\sqrt{t}}, m', r, d). \tag{27}$$

After considering the indirect measurements, suppose that someone measures directly $M^{\text{direct}} = \sum_{j=m+1}^{2m}\Sigma^{(j)}$ on the second group of m subsystems. The empirical average outcome is given by

$$\langle\Sigma\rangle_{emp}^{(m)} = \sum_x \gamma_x Q(x) \tag{28}$$

where Q is the frequency distribution on the alphabet of Σ (similarly as Q_i is the frequency distribution of alphabet $\mathcal{A}_i$ of $\mathcal{L}_i$), and γ_x are some real numbers. In a way similar to the indirect case (but much easier here) we show that the empirical direct average is close to Eq. (19):

$$\mathcal{P}(|\langle\Sigma\rangle_\sigma - \langle\Sigma\rangle_{emp}^{(m)}| > \delta) \leq e(\frac{\delta}{||\Sigma||_{HS}}, m, r, d). \tag{29}$$

From the inequalities (27), (29) we obtain

Lemma 5 *For the measurements on the state $\varrho_{2m,r}^{(\sigma)}$ considered above we have:*

$$\begin{aligned} & \mathcal{P}(|\langle\Sigma\rangle_{emp}^{(m)} - \langle\Sigma^{ind}\rangle_{emp}^{(m)}| > 2\delta) \\ \leq \;& t \cdot e(\frac{\delta}{||\Sigma||_{HS}\sqrt{t}}, m', r, d) + e(\frac{\delta}{||\Sigma||_{HS}}, m, r, d) \\ \leq \;& (t+1)e(\frac{\delta}{||\Sigma||_{HS}\sqrt{t}}, m', r, d) \end{aligned} \tag{30}$$

Proof .- Here triangle inequality and union bound to inequalities (27), (29) suffices together with the properties of $e(\delta, n, r, d)$.

1.3.2 Passing from $\varrho_{2m,r}^{(\sigma)}$-s to their integrals and then to a close-by state

Note that both integration and the measurement of a state to produce the classical distribution of the outcomes are both linear, completely positive, and trace-preserving maps. Thus, Lemma 5 still holds under the replacement $\varrho_{2m,r}^{(\sigma)} \to \int \varrho_{2m,r}^{(\sigma)} d\mu(\sigma)$. Furthermore, if

$$||\varrho_{2m} - \int \varrho_{2m,r}^{(\sigma)} d\mu(\sigma)|| \leq \epsilon \tag{31}$$

and since the trace distance is nonincreasing under the measurement (a TCP map), the output distribution is different by no more than ϵ. In this way we have proven

Lemma 6 *For a state ϱ_{2m} of $2m$ systems satisfying $||\varrho_{2m} - \int \varrho_{2m,r}^{(\sigma)} d\mu(\sigma)|| \leq \epsilon$ we have*

$$\begin{aligned} & \mathcal{P}'(|\langle\Sigma\rangle_{emp}^{(m)} - \langle\Sigma^{ind}\rangle_{emp}^{(m)}| > 2\delta) \\ \leq \;& (t+1)e(\frac{\delta}{||\Sigma||_{HS}\sqrt{t}}, m', r, d) + \epsilon. \end{aligned} \tag{32}$$

where $\mathcal{P}'$ is the probability distribution on outcomes of measurement $\mathcal{L}_1^{\otimes m'} \otimes \ldots \otimes \mathcal{L}_t^{\otimes m'} \otimes \mathcal{M}^{\otimes m}$ induced by the state ϱ_{2m}.

1.3.3 Inferring direct average on $n+m$ samples of general state ϱ_{2m+n} from indirect measurements on m samples

Now we pass to the general permutationally invariant state ϱ_{2m+n}. We want to relate the distance between $\langle\Sigma^{ind}\rangle_{emp}^{(m)}$, the indirect estimation of Σ obtained via LOCC measurements $\{\mathcal{L}_i\}$ on

m of the systems, and the direct estimation $\langle\Sigma\rangle_{emp}^{(m+n)}$ of Σ, we would obtain via the direct measurement $\mathcal{M}$ on the other $n+m$ systems. We have the following:

Theorem 3 *Consider permutationally invariant state ϱ_{2m+n} on $\mathcal{H}^{\otimes 2m+n}$ and $\dim\mathcal{H} = d$. On this state we perform the measurement $\mathcal{L}_1^{\otimes m'} \otimes \cdots \otimes \mathcal{L}_t^{\otimes m'} \otimes \mathcal{M}^{\otimes m+n}$ which induces the probability measure $\mathcal{P}''$. (Note that $\mathcal{P}'$ from Lemma 6 is simply the marginal of $\mathcal{P}''$.) Then we have*

$$\mathcal{P}''(|\langle\Sigma^{ind}\rangle_{emp}^{(m)} - \langle\Sigma\rangle_{emp}^{(m+n)})| > 3\delta) \leq e_1 + e_2 + e_3 \tag{33}$$

where

$$e_1 = 2e^{-\frac{n(r+1)}{2(2m+n)} + \frac{1}{2}d^{\mathbf{2}}\ln n}, \tag{34}$$

$$e_2 = (t+1)2^{-(\frac{\delta^2}{4t||\Sigma||_{HS}^2} - H(\frac{r}{m'}))m' + d\log(\frac{m'}{2}+1)} \tag{35}$$

and

$$e_3 = de^{-\frac{m\delta^2}{\mathbf{8}d||\Sigma||_{HS}^2}}. \tag{36}$$

where $||\cdot||_{HS}$ is the Hilbert-Schmidt norm.

Proof - The parameters e_1,e_2, e_3 come from the generalised quantum de Finetti theorem, Chernoff bound and sampling proposition respectively.

To start with the proof note that from Theorem 1, item 1 we get that for $\varrho_{2m} = \mathrm{Tr}_n\varrho_{2m+n}$ we have $||\varrho_{2m} - \int \varrho_{2m,r}^{(\sigma)} d\mu(\sigma)|| \leq \epsilon$ with $\epsilon = e_1$. Applying then Lemma 6 we get that

$$\mathcal{P}'(|\langle\Sigma\rangle_{emp}^{(m)} - \langle\Sigma^{ind}\rangle_{emp}^{(m)}| > 2\delta) \leq e_1 + e_2 \tag{37}$$

Now we need to connect $\langle\Sigma^{ind}\rangle_{emp}^{(m)}$ with $\langle\Sigma\rangle_{emp}^{(m+n)}$. For this we need sampling Proposition 1 $\mathcal{P}''(||Q_\Sigma^m - Q_\Sigma^{m+n}|| > \delta) \leq de^{-m\delta^2/8d}$ where Q_σ^m is the frequency distribution on outputs of M induced by the state ρ_m (partial trace of ρ_{2m+n} over $m+n$ systems and Q_σ^{m+n} is frequency distribution induced on outcomes of M by state ρ_{m+n} (partial trace of ρ_{2m+n} over m systems)

and d is the dimension of elementary Hilbert space $\mathcal{H}$ (thus ϱ_{2m} is defined on $\mathcal{H}^{\otimes 2m}$. Using Lemma 4 we go to the averages

$$\mathcal{P}''(|\langle\Sigma\rangle_{emp}^{(m)} - \langle\Sigma\rangle_{emp}^{(m+n)}| > 3\delta) \leq e_3. \tag{38}$$

Applying the union bound to Eqs. (37) and (38) we get finally the statement of the theorem.

References and Notes

1. C. H. Bennett, D. P. DiVincenzo, J. Smolin, W. K. Wootters, *Phys. Rev. A* **54**, 3824 (1997).

2. C. Bennett, G. Brassard, *Proc. of IEEE Conference on Computers, Systems and Signal Processing* (1984), pp. 175–179.

3. A. Ekert, *Phys. Rev. Lett* **67**, 661 (1991).

4. K. Horodecki, M. Horodecki, P. Horodecki, J. Oppenheim, *Phys. Rev. Lett.* **94**, 160502 (2005).

5. P. Horodecki, M. Horodecki, R. Horodecki, *J. Mod. Optics* **47**, 347 (2000).

6. D. DiVincenzo, T. Mor, P. Shor, J. Smolin, B. Terhal, *Comm. Math. Phys.* **238**, 379 (2003).

7. K. Horodecki, D. Leung, H.-K. Lo, J. Oppenheim, *Phys. Rev. Lett.* **96**, 081302 (2006).

8. K. Horodecki, M. Horodecki, P. Horodecki, D. Leung, J. Oppenheim (2006). quant-ph/0608195.

9. P. Shor, J. Preskill, *Phys. Rev. Lett.* **85**, 440 (2000).

10. H.-K. Lo, H. F. Chau, *Science* **283**, 2050 (1999).

11. K. Horodecki, M. Horodecki, P. Horodecki, J. Oppenheim, General paradigm for distilling classical key from quantum states. quant-ph/0506189.

12. K. Horodecki, L. Pankowski, M. Horodecki, P. Horodecki, Low dimensional bound entanglement with one-way distillable cryptographic key. quant-ph/0506203.

13. M. Horodecki, P. Horodecki, R. Horodecki, *Phys. Rev. Lett* **80**, 5239 (1998).

14. C. Bennett, G. Brassard, C. Crépeau, U. Maurer, *IEEE Trans. Inf. Th.* **41**, 1915 (1995).

15. P. Shor, J. Preskill, *Phys. Rev. Lett.* **85**, 441 (2000).

16. G. Gottesman, H.-K. Lo, *IEEE Transactions on Information Theory* **49**, 457 (2003).

17. R. Renner, Ph.D. thesis, ETH, Zurich (2005).

18. M. Ben-Or, M. Horodecki, D. W. Leung, D. Mayers, J. Oppenheim, The universal composable security of quantum key distribution (2004).

19. R. Renner, R. Koenig. quant-ph/0410229.